NORTH-HOLLAND

Matrix Bruhat Decompositions With a Remark on the QR (GR) Algorithm

E. Tyrtyshnikov*

Institute of Numerical Mathematics
Russian Academy of Sciences
Leninskij Prospekt, 32-A
Moscow 117334, Russia

Submitted by Richard A. Brualdi

ABSTRACT

In a simple and systematic way we present matrix Bruhat decompositions of two kinds: basic and modified. We show that it is the modified Bruhat decomposition that governs the eigenvalue disorder in the QR (GR) algorithm. This paper can be considered as a commentary on a previous observation about the QR algorithm made by Wilkinson. © *Elsevier Science Inc., 1997*

1. INTRODUCTION

We present a curious relation between the Bruhat decomposition and the QR (GR) algorithm, which seems to have been so far unnoticed.

The Bruhat decomposition is not often discussed in the literature on matrix analysis and numerical methods. Usually one can find it in Lie group theory publications, but prior to this one has to learn very many auxiliary concepts. In Section 2 we give an elementary derivation of principal facts which go with the matrix Bruhat decomposition.

In Section 3 we describe the role of the Bruhat decomposition in the theory of GR (QR) algorithms. However, we should say that this role is of

* E-mail: `tee@inm.ras.ru`.

LINEAR ALGEBRA AND ITS APPLICATIONS 250:61–68 (1997)

0024-3795/97/$17.00
SSDI 0024-3795(95)00453-X

interest only for theory. It has little, if any, value for practice. It can be argued that one could hardly succeed in displaying this role numerically.

The last statement will be explained after reviewing convergence properties of the QR algorithm, a particular case of the GR algorithm.

Given $A_0 = A \in C^{n \times n}$, the GR iterations with no shifts can be written as follows:

$$A_0 = G_1 R_1, \quad A_1 = R_1 G_1; \dots; \qquad A_{k-1} = G_k R_k, \quad A_k = R_k G_k; \dots . \tag{1.1}$$

Here R_k are upper triangular matrices. The matrices G_k should be invertible, and the condition number of the products $Z_k \equiv G_1 \cdots G_k$ should be uniformly bounded in k (see [7]). This demand is obviously met when $G_k = Q_k$ are unitary; in this case we obtain the usual QR algorithm [1, 5]. Well-known theorems on the convergence of the QR algorithm state that under certain hypotheses A_k "essentially" tends to an upper triangular matrix (in the general case, to an upper block triangular matrix).

In this paper we consider a simplyfing assumption which is typical for a first approach to the QR algorithm. Specifically, let A be a nonsingular matrix with eigenvalues of distinct modulus:

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0. \tag{1.2}$$

This means that we can write

$$A = X \Lambda X^{-1}, \qquad \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n), \tag{1.3}$$

where $X \in C^{n \times n}$ is the eigenvector matrix of A.

We shall rely on a nice elementary convergence proof proposed by Wilkinson [8]. In this proof one more simplifying assumption, that X^{-1} has nonzero leading minors, is important. Thanks to this assumption, we may consider an LU decomposition

$$X^{-1} = LU, \tag{1.4}$$

where L is a lower triangular matrix with units on the main diagonal, while U is a nonsingular upper triangular matrix. If (1.2)–(1.4) hold true, then

$$\{A_k\}_{ij} \to 0 \qquad \text{for} \quad i > j, \tag{1.5}$$

$$\operatorname{diag} A_k \to \Lambda. \tag{1.6}$$

In [8] Wilkinson noted that the lack of the LU decomposition for X^{-1} implies that the approximations to eigenvalues are disordered in the limiting matrix for diag A_k. We will show in Section 3 that the Bruhat decomposition of X^{-1} (which always exists) fully specifies that disorder. This result can be regarded as a comment on Wilkinson's observation.

Note that even if the LU decomposition does not exist for X^{-1}, it is guaranteed to exist for arbitrarily small perturbations of X^{-1}. That is why the property (1.4) can be thought of as always fulfilled from the standpoint of real-life computations. Therefore, our specification of the eigenvalue disorder is useless for practice. In addition, the GR algorithm is rarely applied with no shifts, and shifts make the disorder difficult to predict.

2. THE BRUHAT DECOMPOSITION

In contrast to the Bruhat decomposition that figures in group theory (see [3]), its matrix counterpart can be introduced in an elementary way. Assume that $A \in C^{n \times n}$ is nonsingular. Then the Bruhat decomposition of A is defined as

$$A = L_1 \Pi L_1, \tag{2.1}$$

where Π is a permutation matrix, and L_1 and L_2 are nonsingular lower triangular matrices. We also take up another decomposition:

$$A = LPU, \tag{2.2}$$

where P is a permutation matrix, L is a nonsingular lower triangular matrix, and U is a nonsingular upper triangular matrix. Equations (2.1) and (2.2) will be called the basic and modified Bruhat decompositions, respectively.

In numerical linear algebra we are used to dealing with decompositions $A = \bar{P}\bar{L}\bar{U}$ or $A = \tilde{L}\tilde{U}\tilde{P}$ (see [2]), where P, L, U stand for permutation, lower triangular, and upper triangular matrices. In these decompositions $\bar{P}$ and $\tilde{P}$ are *not unique*. In contrast, P and Π from the Bruhat decompositions are placed *in between* triangular factors. In this position they are unique.

Below we give a short proof of the existence of the basic and the modified Bruhat decompositions. We show that the permutation matrices are uniquely determined. We also examine the relation between the basic and the modified decompositions.

THEOREM 2.1. *The modified Bruhat decomposition* (2.2) *exists for any nonsingular matrix* $A \in C^{n\times n}$, *and the permutation matrix* P *is determined uniquely.*

Proof. We show that A can be reduced to a permutation matrix by a sequence of pre- and postmultiplications using appropriate lower and upper triangular matrices. Indeed, consider the first nonzero entry in the first row of A (such an entry exists, since A is nonsingular). Then postmultiplication by an upper triangular matrix can be used to kill all subsequent entries in the first row, and make the pivoting entry equal to 1. Next, premultiplication by a lower triangular matrix can be used to zero all entries which are located below the pivoting one in its column. Once this is done, we find a new pivoting entry, that is, the first nonzero entry in the second row of the current matrix. Using postmultiplication, we annihilate all entries to the right of the pivoting one in the second row, and using premultiplication, then we get rid of all entries below the pivoting one in its column. And so on. Since A is nonsingular, the above process produces a permutation matrix after $n-1$ steps. Let $\hat{P} = \hat{L}A\hat{U}$ be the resulting permutation matrix, where $\hat{L}$ and $\hat{U}$ are the products of involved elementary lower (for premultiplication) and upper (for postmultiplication) triangular matrices. Evidently, setting $P = \hat{P}$, $\hat{L} = \hat{L}^{-1}$, $U = \hat{U}^{-1}$, we arrive at (2.2).

Write $\hat{P} = [\delta_{\sigma(i),j}]_{i,j=1}^{n}$, where δ is the Kronecker symbol ($\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ij} = 1$ for $i = j$). For any matrix A we may consider the ith row and its leading subrows (below we adopt Matlab notation):

$$r_i(j) = A(i, 1:j), \quad j = 1, \dots, n.$$

Then let $\sigma(A; i)$ be the minimal j such that

$$r_i(j) \notin \operatorname{span}\{r_1(j), \dots, r_{i-1}(j)\}. \tag{2.3}$$

Obviously, $\sigma(i) = \sigma(\hat{P}; i)$. It is easy to check that the property (2.3) is preserved whenever A is postmultiplied by a nonsingular upper triangular matrix or premultiplied by a nonsingular lower triangular matrix. Hence,

$$\sigma(i) = \sigma(\hat{L}A\hat{U}; i) = \sigma(A; i).$$

We see that $\sigma(A; i)$ is uniquely determined for A, and thus $\hat{P}$ is uniquely determined. This completes the proof. ∎

REMARK. The triangular factors L and U are not unique. For instance, set

$$J = \begin{bmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{bmatrix}_{n \times n}. \tag{2.4}$$

Then for any lower triangular matrix L, the matrix $U = JLJ$ is upper triangular and

$$A \equiv LJU = L^2 J.$$

We thus have two different Bruhat decompositions for A (of course, with the same permutation matrix).

The matrix J will help us to establish a relation between the basic and the modified Bruhat decompositions. If A is nonsingular, then AJ is also nonsingular, and hence by Theorem 2.1 we may write

$$AJ = LPU. \tag{2.5}$$

This is the modified Bruhat decomposition for AJ, and it immediately follows that

$$A = L(PJ)(JUJ), \tag{2.6}$$

which is the basic Bruhat decomposition for A. On the other hand, a basic Bruhat decomposition

$$AJ = L_1 \Pi L_2 \tag{2.7}$$

can be readily transformed to a modified Bruhat decomposition

$$A = L_1(\Pi J)(JL_2 J). \tag{2.8}$$

Thus, there holds the following

THEOREM 2.2. *The basic Bruhat decomposition* (2.1) *exists for any nonsingular $A \in C^{n \times n}$, and the permutation matrix Π is uniquely determined. If we write $\Pi = \Pi(A)$ in the basic Bruhat decomposition and $P = P(A)$ in the modified Bruhat decomposition, then*

$$\Pi(A) = P(AJ)J, \tag{2.9}$$

$$P(A) = \Pi(AJ)J. \tag{2.10}$$

3. EIGENVALUE DISORDER IN THE GR ALGORITHM

We first recall two principal relationships of the GR iterations (see [6, 8, 9] for the case of the QR algorithm):

$$A_k = Z_k^{-1} A Z_k, \qquad Z_k \equiv G_1 \cdots G_k, \tag{3.1}$$

$$A^k = Z_k U_k, \qquad U_k \equiv R_k \cdots R_1. \tag{3.2}$$

These relationships can be easily derived from (1.1).

We assume that (1.2) and (1.3) hold true. However, instead of (1.4) we consider here a modified Bruhat decomposition

$$X^{-1} = LPU. \tag{3.3}$$

We thus dispense with the previous hypothesis, that X^{-1} has nonzero leading minors.

It follows from (3.2) that $Z_k = A^k U_k^{-1}$ and consequently

$$A_k = \left(U_k A^{-k}\right) A \left(A^k U_k^{-1}\right). \tag{3.4}$$

Allowing for (1.3), we find

$$A^k U_k^{-1} = X \Lambda^k (LPU) U_k^{-1} \tag{3.5}$$

and then

$$A_k = \Phi_k^{-1} \left[P^{-1} \Lambda^k (L^{-1} \Lambda L) \Lambda^{-k} P \right] \Phi_k, \tag{3.6}$$

where

$$\Phi_k = P^{-1} \Lambda^k P U U_k^{-1}. \tag{3.7}$$

The matrices Φ_k and Φ_k^{-1} are upper triangular. Since $\operatorname{cond}_2 Z_k$ is uniformly bounded in k, these matrices remain bounded as $k \to \infty$. To prove this we write

$$\begin{aligned} \Phi_k &= P^{-1} (\Lambda^k L^{-1} \Lambda^{-k}) \Lambda^k (LPU) U_k^{-1} \\ &= P^{-1} (\Lambda^k L^{-1} \Lambda^{-k}) X^{-1} Z_k \end{aligned}$$

and taken into account that $\Lambda^k L^{-1} \Lambda^{-k} \to I$ so long as the assumption (1.2) is fulfilled.

Due to (1.2) and since $L^{-1}\Lambda L$ is a lower triangular matrix, we get

$$\Lambda^k (L^{-1}\Lambda L)\Lambda^{-k} \to \operatorname{diag}(L^{-1}\Lambda L) = \Lambda. \tag{3.8}$$

Therefore, the matrix in the square brackets in (3.6) tends to $P^{-1}\Lambda P$ with $k \to \infty$. Thus, as $k \to \infty$ we have

$$\operatorname{diag} A_k \to P^{-1}\Lambda P, \tag{3.9}$$

and at the same time A_k is becoming "essentially" upper triangular, that is,

$$\{A_k\}_{ij} \to 0 \qquad \text{for} \quad i > j. \tag{3.10}$$

Thus we have proved the following

THEOREM 3.1. *Assume that $A \in C^{n \times n}$ is a nonsingular matrix with eigenvalues* (1.2) *of distinct modulus, and the GR algorithm* (1.1) *generates a sequence of matrices A_k such that the condition numbers of $Z_k \equiv G_1 \cdots G_k$ are uniformly bounded in k. Then the relations* (3.9) *and* (3.10) *hold true, where the permutation matrix P is determined by the modified Bruhat decomposition of X^{-1} from the spectral decomposition* (1.3) *of A.*

This result admits a generalization to the case when A has eigenvalues of equal modulus. In this case, one has to prove that the matrices A_k "essentially" converge to a block upper triangular matrix. The columns of X now determine a Jordan basis for A, and again the Bruhat decomposition of X^{-1} sheds light on the disorder of blocks in the limiting matrix.

I would like to thank Gilbert Strang and David Watkins for their helpful comments on the manuscript.

REFERENCES

1 J. G. F. Francis, The QR transformation: A unitary analogue to the LR transformation, Parts I, II, *Comput. J.* 4:265–272, 332–345 (1961).

2 G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins U.P., 1989.

3 D. Yu. Grigor'ev, Lower bounds in the algebraic computational complexity (in Russian), *Zap. Nauchn. Sem. LOMI* 118:25–82 (1982).
4 Kh. Ikramov, *Nonsymmetric Eigenvalue Problem* (in Russian), Nauka, 1991.
5 V. N. Kublanovskaya, On some algorithms of the complete eigenvalue problem, *USSR Comput. Math. Phys.* 3:637–657 (1961).
6 V. V. Voevodin, *Computational Fundamentals of Linear Algebra* (in Russian), Nauka, 1977.
7 D. S. Watkins and L. Elsner, Convergence of algorithms of decomposition type for the eigenvalue problem, *Linear Algebra Appl.* 143:19–47 (1991).
8 J. H. Wilkinson, Convergence of the LR, QR, and related algorithms, *Comput. J.* 8(1):77–84 (1965).
9 J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.

Received 6 January 1995; final manuscript accepted 4 May 1995